

*Spatiotemporal Analyses of Traffic Flow Relationships*  
KUHN, Kenneth; NICHOLSON, Alan

# SPATIOTEMPORAL ANALYSES OF TRAFFIC FLOW RELATIONSHIPS

*Kenneth Kuhn; University of Canterbury, Christchurch, New Zealand;*  
[kenneth.kuhn@canterbury.ac.nz](mailto:kenneth.kuhn@canterbury.ac.nz)

*Alan Nicholson; University of Canterbury, Christchurch, New Zealand;*  
[alan.nicholson@canterbury.ac.nz](mailto:alan.nicholson@canterbury.ac.nz)

## ABSTRACT

Empirical data from the Tokyo Metropolitan Expressway show how relations between traffic flow, density and speed evolve as the data are aggregated spatially and temporally. Considering larger geographic areas or longer time frames consistently reduces scatter associated with fundamental traffic flow relationships, according to quantitative results presented here. We argue that econometric results largely explain the somewhat counter-intuitive finding. The relatively strong explanatory power of traffic flow relationships in aggregated data stems from correlations between observations of traffic on adjacent links and at adjacent points in time, as well as loop detector observational error. The data actually support a stronger claim that estimates of flow across a large area based on the aggregate fundamental diagram outperform estimates based on the summation of terms from link-specific fundamental diagrams. This result again matches econometric findings.

*Keywords: traffic flow theory, macroscopic fundamental diagram, data aggregation*

## INTRODUCTION

Traffic conditions at specific locations on a roadway network at specific times are often described in terms of vehicle flow rate ( $q$ ), mean vehicle speed/velocity ( $v$ ) and density/concentration of vehicles ( $k$ ). Early work by Greenshields (1935) showed sections of roadway have characteristic linear relationships between speed and density. Lighthill and Whitham (1955) used fluid dynamics to derive a general relation between vehicle flow rate and density. Flow increases with density until a maximum flow is achieved, then decreases until a jam density is reached where flow is zero. This relation agrees with Greenshields work given that flow equals speed multiplied by density ( $q=kv$ ). Greenberg (1959) and Drake et al. (1967) were two of the first of many works refining the Greenshields and Lighthill-Whitham models. Site-specific design characteristics determine what parameter values make sense at different locations.

Data regarding flow, speed, and density recorded at specific locations over short time intervals are frequently plotted to show empirical relationships. Produced figures, like Figure 1 below, generally support hypothesized relationships but contain significant scatter. The data points shown in Figure 1 are based on observations of flow rate and speed over ten minute time windows at section 27 of the Central Circular Route in Tokyo, Japan. (This data will be described in further detail later in this paper.)

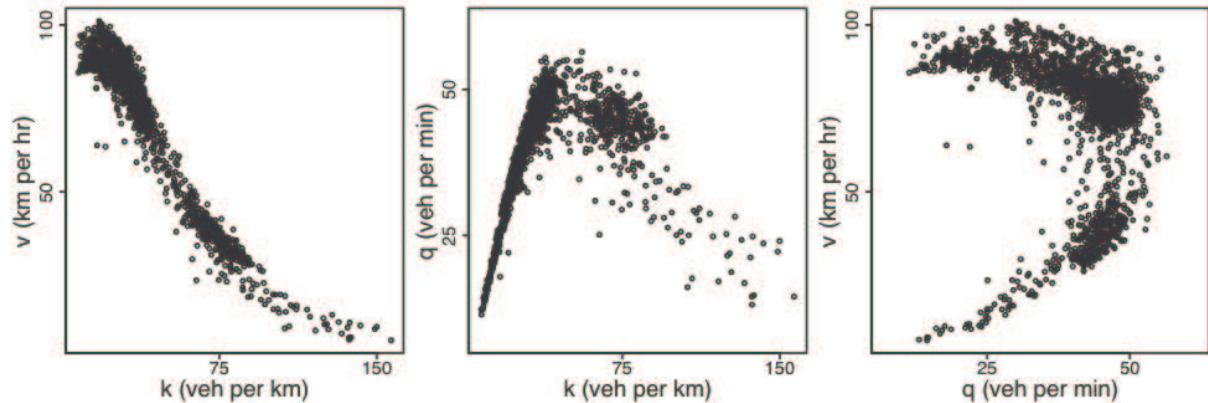


Figure 1: Traffic flow relationships: section 27, Central Circular Route, Tokyo Metropolitan Expressway.

## Urban-area Traffic Analyses

Authors analyzing traffic data aggregated across large urban areas have found relationships between flow, speed and density similar to those found in Figure 1 with comparable or reduced amounts of scatter. These results are surprising given that it is easier to imagine increased vehicle densities causing decreased speeds on a particular section of one road rather than across an entire city. In addition, one would expect different sections of roadway to have vastly different sensitivities to changing traffic conditions.

In an influential early work, Thomson (1967) analyzed empirical data from central London and discovered an inverse relationship between average vehicle speed and the flow of traffic. Flow was calculated averaging measurements of traffic flows on particular sections of the London roadway network weighted according to section length. Each individual data point represented average flow and average speed across an entire year of measurement. Wardrop (1968) extended this work by formalizing a relation between speed, flow, average street width and traffic signal spacing. Godfrey (1969) investigated the relationship between average vehicle speed and the density of vehicles in a city, as well as between speed and vehicle miles travelled within a city. The author found that average speed declined as vehicle density increased and that the maximum vehicle miles travelled occurred when vehicles were travelling at around 10 miles per hour. Godfrey's data included hourly observations of speed and density.

In 1979, Herman and Prigogine discovered that the velocity of moving cars is proportional to the fraction of cars that are moving (as opposed to stopped in traffic) raised to a power, and the fraction of cars stopped is proportional to the density of vehicles raised to a power. In 1984, Herman and Ardekani confirmed these findings using probe vehicles driving around

Austin. Williams et al. (1987) used simulation to find that, in data describing city-wide traffic conditions “the general similarity of the K (density) - V (speed) - Q (flow) patterns to their characteristic counterparts for individual road sections is striking.” Mahmassani et al. (1990) confirmed these results using simulation data. Olszewski et al. (1995) fit models relating area-wide measurements of speed, flow, and density.

In 2007, Daganzo hypothesized that the flow of vehicles exiting an urban *neighbourhood* roughly measured the flow of vehicles within the neighbourhood and was related to the number of vehicles circulating through the neighbourhood. The author noted that control of inflow, the rate of vehicles entering a neighbourhood, could be used to maximize the flow of vehicles within the neighbourhood and outflow from the neighbourhood. Daganzo (2007) offered the belief that neighbourhoods should be defined “of dimensions comparable with a trip length.” Geroliminis and Daganzo (2008) then used data from fixed detectors and probe vehicles in a 10  $km^2$  neighbourhood in downtown Yokohama to show a linear relation between an urban neighbourhood’s travel production, flow multiplied by network length, and the neighbourhood’s outflow. Results were offered concerning *weighted flow*, based on the weighted averaging technique of Thomson (1967), and *unweighted flow*, which was the simple average of flows reported on different detectors. Geroliminis and Daganzo (2008) also showed a relation between travel production and vehicle accumulation, density multiplied by network length, in urban neighbourhoods. Making a stronger claim than earlier works, the authors wrote that this relationship between flow and density was actually sharper for an entire urban area than for individual sections of roadway (“the scatter nearly disappeared and points grouped neatly along a smoothly declining curve”). In this publication, the authors neither specified an appropriate functional form for an area-wide flow-density relationship nor attempted to quantitatively describe scatter. Rather, figures provide visual evidence of relatively sharp (noisy) speed-flow-density relationships for urban-area (link-specific) data. In a related publication (Daganzo and Geroliminis, 2008), the same authors do describe how area-wide models of relationships between traffic flow variables can be constructed based on link-specific analyses and/or assumptions of homogeneity. None of the articles surveyed above considered the impact of temporal aggregation of data on speed-flow-density relationships.

## **Econometric Literature**

A number of works in econometrics examine the impact of data aggregation. Yule (1903) warns of Simpson's Paradox, the idea that small- and large-scale relations may conflict in cases involving confounding variables. Thiel (1954) showed that aggregating multiple models that perfectly describe micro-economic relations can lead to inaccurate macro-economic models due to *aggregation bias*.

Grunfeld and Grilches (1960) noted alternately that macro-economic models often outperform aggregated micro-economic models if the smaller-scale models are imperfect. Grunfeld and Grilches wrote that in empirical data “it is quite likely that a macro equation will have a higher  $R^2$  than a micro equation.” A reader familiar with transportation engineering may be reminded of trip generation, where techniques based on modelling how many trips

are generated by *zones* generally have higher coefficients of determination than techniques based on modelling how many trips are generated by *households*. According to Grunfeld and Grilches (1960), larger-scale relations will appear sharper than smaller-scale relations whenever small-scale residual errors are less highly correlated than variables of interest. Mathematical explanations of this result have been provided by Grunfeld and Grilches (1960), Aigner and Goldfeld (1974) and Shin (1987). In a related manner, Grunfeld and Grilches (1960) also noted that noise in micro data can lead to stronger relations between macro variables. A full exposition of this point was made by Aigner and Goldfeld (1974). Stronger still, Grunfeld and Grilches (1960) point out that macro-level dependent variables are often better predicted by macro-level independent variables directly rather than by aggregating results from several micro-level relationships. Granger (1987) later noted that aggregate data values are often largely determined by factors common to micro-level data. In such cases macro-level models can dramatically outperform models based on summing terms from micro-level relationships (Ibid.).

Mathematically, the econometric results described above are largely based on analyses of ordinary least squares linear regression. As such, the results are of clear relevance to many traffic flow modelling efforts but less relevance to others (such as the parametric model proposed by Daganzo and Geroliminis in 2008). We are investigating extending the econometric literature to consider more precisely different mathematical relationships used in traffic flow theory. As is, the econometric literature is still relevant. In the context of traffic flow theory, demand for travel is a common factor of the type Granger described. Flow, density and speed observations taken at the same time at different links on a roadway network are often significantly correlated due both to common dependence on demand for travel and to functional relationships (Munakata, 2007; Nicholson and Munakata, 2009; Nicholson et al., 2010). In addition, there is significant noise in link-specific observations of traffic flow, speed, and density. Thus the econometric literature largely explains the results of urban-area traffic analyses.

## **DATA ANALYSES**

This study is based on data collected on an 11.9 *km* section of the Central Circular Route, which is part of the Tokyo Metropolitan Expressway system. The study area is shown in Figure 2 and covers the area between Kasai and Horikiri junctions. Data came from loop detectors spread across all north-bound lanes of traffic at 36 locations roughly 300 metres apart. Flow rate and average speed were recorded minute-by-minute at each of the locations for 93 days from October 2006 to January 2007. This data set is remarkable for its detail, the volume of data collected and the prevalence of high density conditions. Further details regarding the data are available in Munakata, 2007.



Figure 2: Study area.

Since the data available includes flow and speed estimates, work here focuses on the relationship between the two. Aggregating data spatially and temporally, we do not divide our study area into links and assume that data at specific locations describes traffic flow throughout these links. Instead we assume only that data at a specific location describes conditions at that location, and aggregate data over the set of locations where data are available. We use the definitions of unweighted flow rate and velocity found in Geroliminis and Daganzo (2008). Assume  $q_{lt}$  and  $v_{lt}$  are the flow rate and speed reported at location  $l$  at time  $t$ , respectively. Then the aggregate flow rate across locations from  $l_1$  to  $l_2$  over the time period from  $t_1$  to  $t_2$  is defined as follows.

$$q_{AGG} = \frac{\sum_{l=l_1}^{l_2} \sum_{t=t_1}^{t_2} q_{lt}}{(l_2 - l_1)(t_2 - t_1)} \quad (1)$$

Note that we take the average of the observed flow rates. This formulation will be discussed further later.

Average speed is more difficult to calculate as different measurements from different locations and/or different points in time are based on different numbers of observed cars. Here we note that density, like flow, can be averaged across detector stations given the reasonable assumption that detector stations are roughly equivalent in area. Average speed is defined as flow divided by density, given the accepted equation  $q = kv$ .

$$v_{AGG} = \frac{q_{AGG}}{k_{AGG}} = \frac{\sum_{l=l_1}^{l_2} \sum_{t=t_1}^{t_2} q_{lt}}{\sum_{l=l_1}^{l_2} \sum_{t=t_1}^{t_2} (q_{lt} / v_{lt})} \quad (2)$$

Here, we seek to quantitatively investigate how strong the relationships between speed and flow are for different data sets. In order to do this, we decided to fit various models to the

data and use residual error terms to characterize the strengths of the relationships. The larger the absolute values of the residual error terms are, the weaker the relationships. There is an ongoing debate regarding the appropriate forms to use when modelling speed-flow-density relationships (see, for instance, Papageorgiou, 1998). We wish to side-step this debate as much as possible. Much of the past research considering area-wide traffic relationships was based on visual inspection of generated plots (Williams et al. 1987; Geroliminis and Daganzo, 2008; etc.). We aim to capture the spirit of the analyses via visual inspection, and look only at the relationship between data points comprised of concurrent observations of speed and flow.

## **General Forms of Speed-Flow Relationships**

The modelling of relationships among traffic variables began when Greenshields claimed speed declined linearly as density increased. Given that flow equals speed multiplied by density, then flow is a quadratic function of speed as follows.

$$q = \phi v^2 + \psi v \quad (3)$$

We expect  $\phi$  to be negative and have a significantly smaller absolute value than the positive term  $\psi$ . This matches what we'd expect given that  $q = kv$  yields the following relationship between density and speed.

$$v = (-\psi / \phi) + (1 / \phi)k \quad (4)$$

In this study, we largely ignore density but seek to find values for  $\phi$  and  $\psi$  that fit data on flow and speed. For a given data set (at any level of spatial and/or temporal aggregation) we select estimates of  $\phi$  and  $\psi$  modelling flow as a function of speed, as in equation (3).

It is now widely recognized that Greenshields' model of traffic describes relationships between variables well for such a simple model but falls short in a few areas. At the lowest and highest observable densities, the linear relationship between speed and density breaks down (see Figure 1). The relationship between flow and speed is not symmetric as suggested by Greenshields' work (see Figure 1 again). There have been several efforts to refine Greenshields work or devise alternate models, including the works of Greenberg (1959) and Drake et al. (1967) cited earlier.

Drake et al. (1967) suggested a non-linear speed-density relationship that (in generalized form) yields the speed-flow relationship model form of equation (5). The authors of that study used visual inspection of 21 graphs, as well as analyses of coefficients of determination and standard errors, to suggest this model form over numerous alternatives.

$$q = \alpha(v + \beta \ln(v))^\chi \quad (5)$$

Importantly, the Drake model form shown in equation (5) was also shown, via visual inspection, to match area-wide traffic flow and speed data better than alternative model forms in two subsequent studies (Williams et al., 1987; Olszewski et al., 1995).

There have been significant research efforts since the publication of the Drake model suggesting more sophisticated models. The work of Daganzo and Geroliminis (2008) is notable as it specifically concerns area-wide traffic flow modelling. The exact methodology outlined in the cited work would have been difficult to apply in the context of the computational study presented here. The methodology required an assumption of homogeneity among links that we were hesitant to make in the context of this study or knowledge of several parameters describing individual links of the roadway network being surveyed (such as the speed of backwards moving traffic waves). One important contribution of the Daganzo and Geroliminis work is that the constraints different densities impose on flow rate can be considered separately and then used to provide an overall picture of the flow-density relationship.

Here, we use local regression (also known as locally weighted scatterplot smoothing) of empirical data to develop models of fundamental traffic flow relationships that appear similar to those proposed by Daganzo and Geroliminis. Local regression is able to capture asymmetries in speed-flow relationships. More generally, this nonparametric method is relatively less sensitive to changes in the shape of the fundamental traffic relationships than either the Greenshields or Drake model forms described above. It also makes clear intuitive sense that the magnitudes of residuals of scatterplot smoothing efforts be used to describe the amounts of scatter in graphs. When performing scatterplot smoothing, two parameters must be set: the degree of local polynomials considered, and the smoother span ( $f$ ) which gives the proportion of all data points considered when local polynomials are fit. Polynomials of degree 2 were used here, based on visual inspection of speed-flow relationships (such as in Figures 1 and 3) and consideration of Greenshields' model. A range of values for the smoother span was considered, with  $f$  varyingly set to 0.15, 0.25, and 0.4.

## **Scatter**

The first analyses undertaken sought to compare the scatter in the relationships between speed and flow for different data sets. The goal was to examine the suggestion of previous studies that scatter is reduced as data are aggregated. Figure 3 shows examples of two plots generated during this analysis, with a relatively small and a relatively large amount of scatter. The graph at left is based on data aggregated over 15 minute blocks of time across detector stations 5 through 8, while the graph at right is based on minute-by-minute observations of speed and flow at detector station 22.



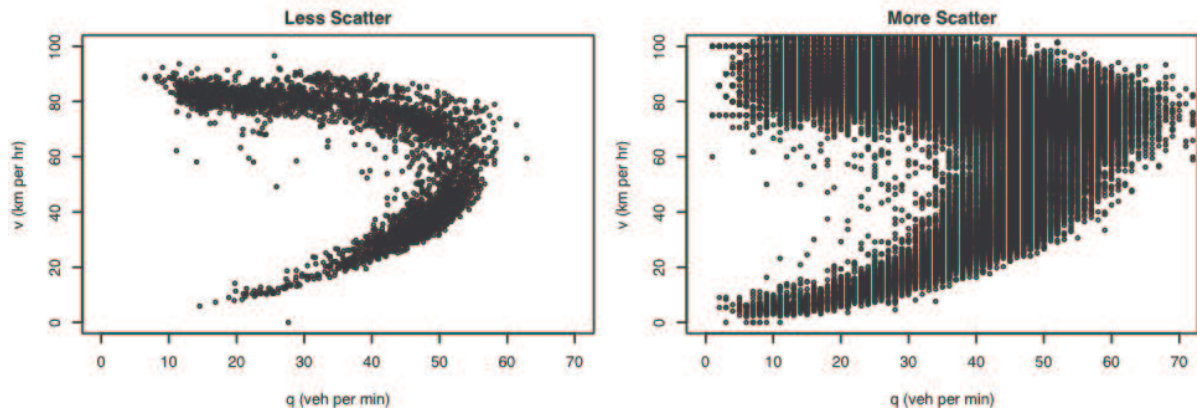


Figure 3: Examples of plots of speed vs flow.

Figure 3 brings up a few interesting points. There are a few data points plotted where speed is close to zero but flow is suspiciously substantial. Such data likely reflect the fact that loop detector data are often imperfect. In estimating scatter, it is worthwhile to exclude extreme outlying data points as these may be caused by data collection errors. The plot containing less scatter contains significantly fewer data points than the plot containing more scatter. As data are aggregated across longer periods of time, the result is fewer available data points. As data are aggregated across larger geographic areas, the result is fewer speed-flow relationships to plot and/or model. Note that estimates of flow and speed are similar in magnitude regardless of spatiotemporal frame of analysis. This stems from the fact that we are looking at the average flow rate per detector station. The result is that it is not necessary to normalize the data to evaluate scatter. Finally note that the differences between relationships containing small and large amounts of scatter are substantial. We expect differences between the magnitudes of residuals to depend more on the data set analyzed and less on the model form used. This is reassuring, as the choice of speed-flow model form is itself an open research question.

In this study, the 36 loop detector stations were variously arranged into 36 groups of 1, 18 groups of 2, 12 groups of 3, 9 groups of 4, 6 groups of 6, 4 groups of 9, 3 groups of 12, 2 groups of 18, and 1 group of 36 adjacent stations. For each group of stations, flow and speed data were collected in blocks of 1, 2, 5, 10, 15, 20, and 30 minutes. For each block of detector stations, Greenshields, Drake and local regression models of speed-flow relationships were found. The residual data values associated with all the speed-flow models at the same level of temporal and spatial data aggregation were then grouped together. As mentioned earlier, different levels of aggregation were associated with different numbers of data points.

The median, 75<sup>th</sup> and 90<sup>th</sup> percentile of the absolute values of residuals were chosen as metrics of scatter. The goal was to see how far typical or somewhat unusual observations of flow rate were from modelled values, but not to consider significantly outlying data points. The analysis of residuals is common in exploratory data analysis. Note that alternate metrics of scatter such as sample standard deviation and range are less robust and would strongly reflect extreme outlying data points likely caused by loop detector operational error. Also note that Figure 3 indicates that the differences between relationships exhibiting small and



large amounts of scatter are substantial and should be reflected by any reasonable metrics of scatter.

Figures 4, 5, and 6 each contains five plots. Each plot is a graphical illustration of the values the various metrics for scatter obtained depending upon level of aggregation in time (x axis) and space (y axis). Relatively darker cells indicate higher residual error terms, and hence more noisy relationships. Figure 4 shows results regarding the medians of the absolute residuals, Figure 5 shows results for the 75<sup>th</sup> percentile of the absolute residuals, and Figure 6 for the 90<sup>th</sup> percentile of the absolute residuals. Separate shading schemes are used for each figure.

It is very clear that aggregating data over larger geographic areas or longer time frames reduces scatter associated with fundamental traffic flow relationships. As in earlier studies, we find that the Drake model form matches empirical data better than the Greenshields model form. Techniques based on local regression produced smaller typical (median) absolute residuals than either the Greenshields or Drake models. There is weak evidence suggesting the local regression results improve as the smoothing span decreases from 0.4 to 0.25 or 0.15. It should be noted that the local regression techniques do not provide closed form equations for speed-flow relationships and thus may be less useful for traffic engineering purposes than Greenshields or Drake models.

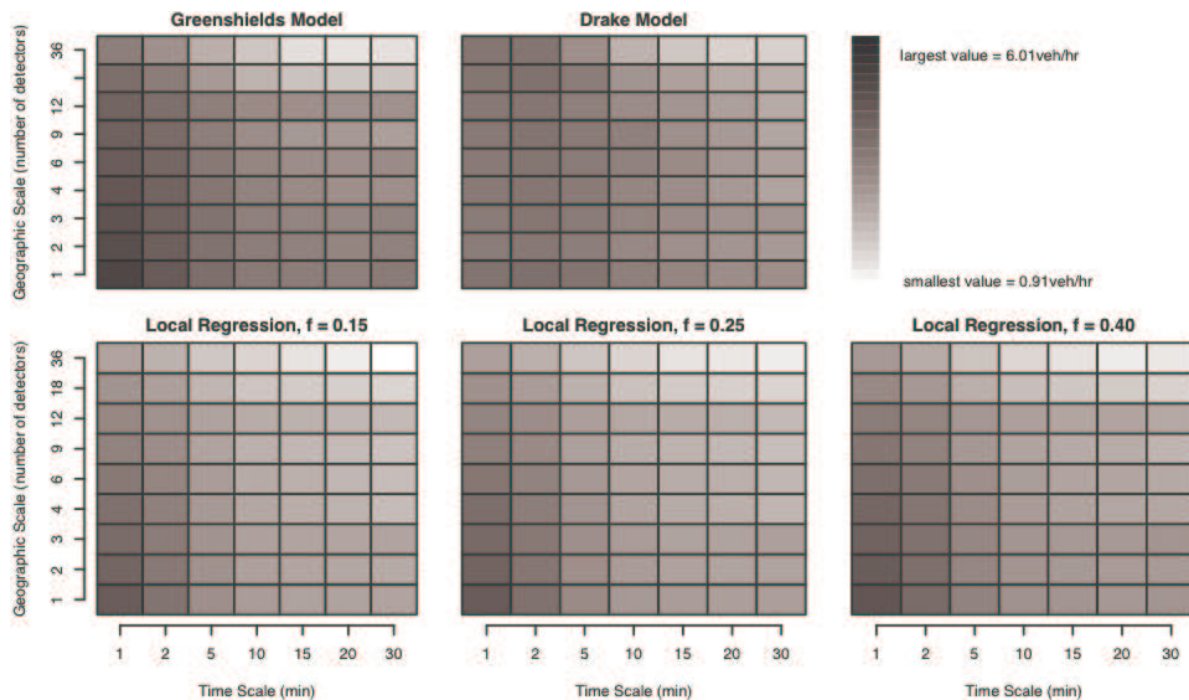


Figure 4: Median absolute residuals as a function of spatial and temporal frame of analysis.

*Spatiotemporal Analyses of Traffic Flow Relationships*  
 KUHN, Kenneth; NICHOLSON, Alan

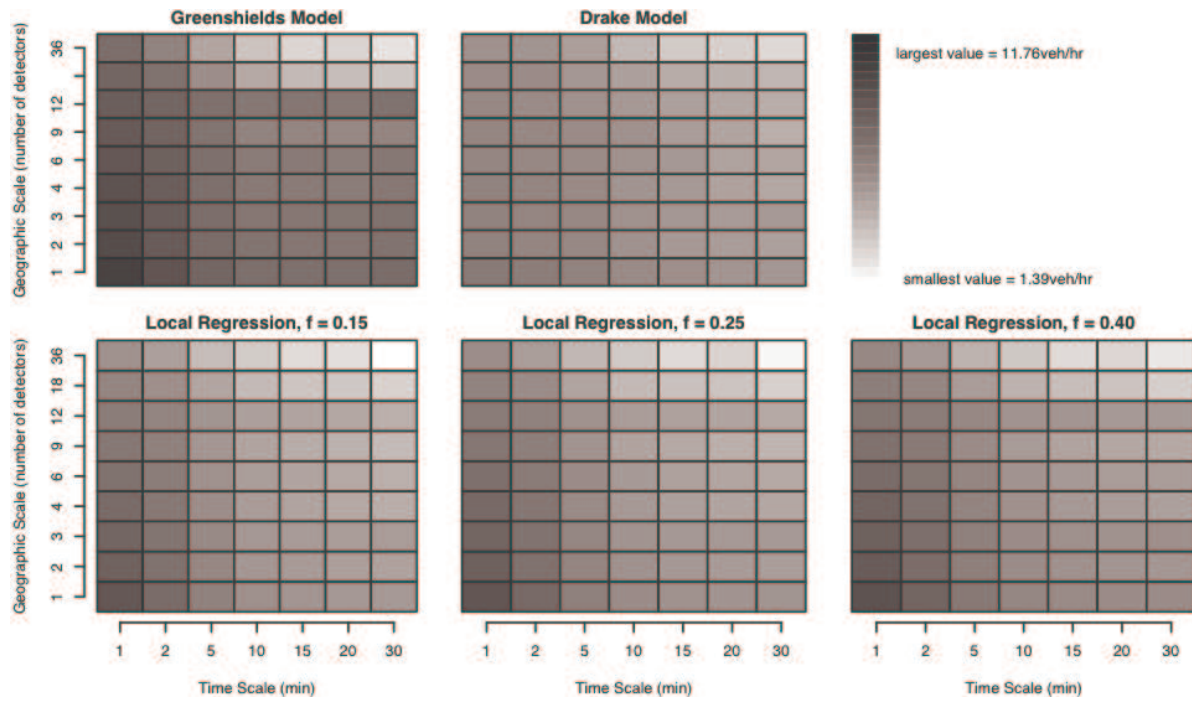


Figure 5: 75<sup>th</sup> Percentile absolute residuals as a function of spatial and temporal frame of analysis.

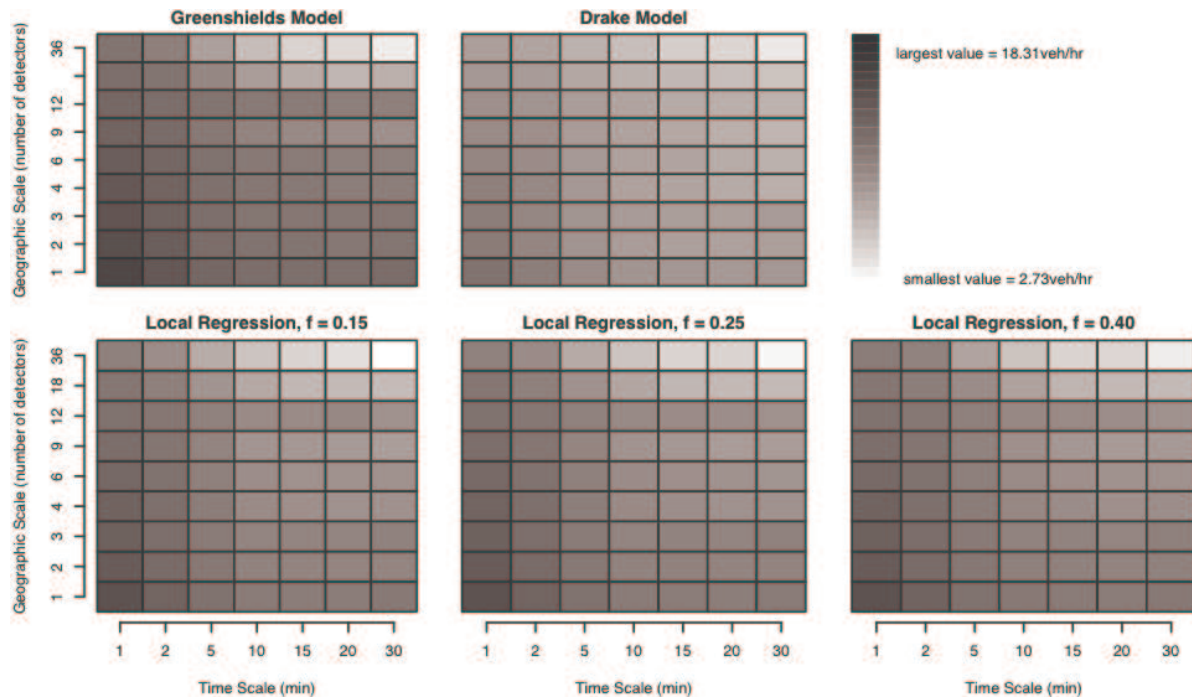


Figure 6: 90<sup>th</sup> Percentile absolute residuals as a function of spatial and temporal frame of analysis.

## Modelling Aggregate Flow

The result regarding scatter can be partially explained by the fact that we are averaging flow rate estimates. An outlying data point recorded at one detector station at one time will have reduced significance as we aggregate increased amounts of data. The next analysis performed examined how well we are able to estimate aggregate flow rate based on aggregate and disaggregate speed-flow relationships. In this analysis, the significance of an outlying data point recorded at one detector station at one time will be minimal regardless of the scale of analysis used.

We seek to model the sum of the flows observed across the 36 links over the course of 10 minutes. This is done first by summing up estimates of the flows on each link based on link-specific observations of average speed and link-specific traffic flow relationships. An alternate approach uses estimates of the average speed across the 36 links and the aggregate traffic flow relationship. The two approaches were applied in conjunction with the three model forms used here, and residual data values were noted. (Only one local regression model, with  $f = 0.25$ , was used for simplicity's sake.) Histograms of obtained residuals are provided in Figure 7.

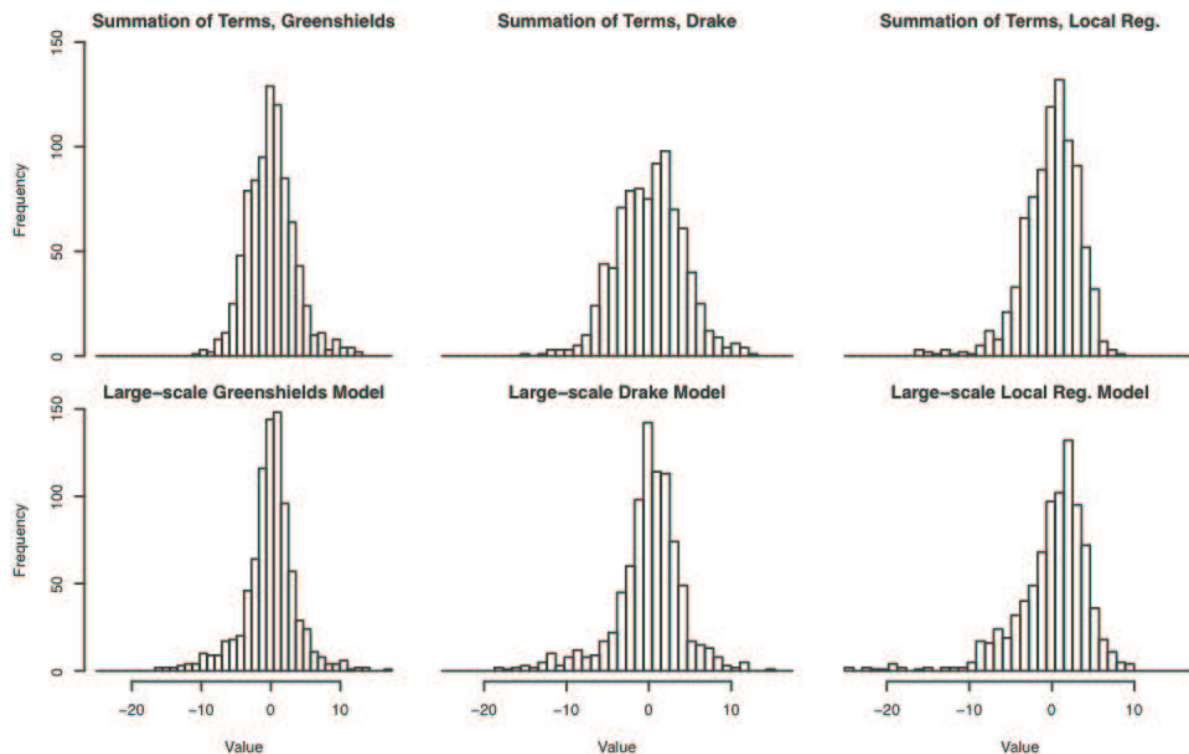


Figure 7: Histograms of residual values for different models of aggregate flow.

Figure 7 shows that the large-scale models are able to predict aggregate traffic flow at least as well as sums of terms from link-specific models. In terms of the metrics used earlier, the medians and 75<sup>th</sup> percentiles of the absolute values of residual errors are all smaller using

aggregate data. This is true despite the fact that the link-specific models use substantially more input data, and are able to capture the different sensitivities of different links to changes in traffic conditions. Interestingly, the Greenshields model appears to be as accurate as the alternatives when modelling large-area traffic flow. Aggregating data over larger areas likely yields fewer observations of traffic conditions characterized by exceptionally low or high levels of density. This would also help explain why large-area models, especially those based on the Greenshields model form, are relatively more accurate than small-scale models.

There is a growing interest in large-scale urban traffic control strategies. Such studies require modelling of large-scale traffic conditions. The results of this work demonstrate that there is not necessarily anything gained by modelling individual links of a roadway system when estimating traffic conditions across a large urban area. Even for smaller scale analyses considering individual trips across multiple sections of a roadway network, the work shown here indicates that it would be best to model conditions across the links of interest all together rather than breaking analysis down by link. This has recently been shown to be true for measures of travel time reliability (Nicholson and Munakata, 2009; Nicholson et al., 2010). The work done here indicates it may also be true for a more basic estimation of average speed (i.e., trip travel time).

The large-scale models of traffic conditions described above were exceedingly simple. The development of more advanced models of aggregate traffic flow relationships is warranted. In particular, further work could create models of large-scale traffic flow variables that capture link-specific sensitivities to changes in traffic conditions while also looking for common factors more important at larger scales of analysis, such as travel demand.

## **CONCLUSION**

Work presented here shows that the fundamental traffic flow relationships between speed, flow and density explain more of the variability in observed data when the data are aggregated over larger geographic areas or longer time frames. This counter-intuitive finding agrees with previous empirical studies including Geroliminis and Daganzo (2008). This work also shows how these results agree with the econometric literature including Grunfeld and Griliches (1960). Data analyzed here supports a stronger claim; large-area traffic flow variables are more accurately modelled using a large-area fundamental diagram as opposed to aggregating link-specific models. This surprising result again matches what Grunfeld and Griliches (1960) found. Correlations between traffic flow variables on different links largely explain the results. Error in the observation of traffic flow variables may also be relevant. The results found here indicate that we should be wary of techniques involving summing terms estimated on a link-by-link basis. Further research is needed to identify how to better model traffic conditions for trips or analyses involving significant geographic areas.

## REFERENCES

- Aigner, D. J. and S. M. Goldfeld (1974). Estimation and prediction from aggregate data when aggregates are measured more accurately than their components. *Econometrica*, 42.
- Daganzo, C. F. (2007). Urban gridlock: macroscopic modeling and mitigation approaches. *Trans. Res. B* 41.
- Daganzo, C. F. and N. Geroliminis (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Trans. Res. B* 42.
- Drake, J. S., Schofer, J. L. and A. D. May (1967). A statistical analysis of speed-density hypotheses. *Proc. Third Int'l. Symp. on Theory of Traf. Flow*.
- Geroliminis, N. and C. F. Daganzo (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Trans. Res. B* 42.
- Godfrey, J. W. (1969). The mechanism of a road network. *Traf. Eng. and Control* 11.
- Granger, C. W. J. (1987). Implications of aggregation with common factors. *Econometric Theory* 3.
- Greenberg, H. (1959). An analysis of traffic flow. *Oper. Res.* 7.
- Greenshields, B. D. (1935). A study of traffic capacity. *Pro. of the Highway Res. Board* 14.
- Grunfeld, Y. and Z. Griliches (1960). Is aggregation necessarily bad?. *Rev. of Econ. And Stats.* 42.
- Herman, R. and S. Ardekani (1984). Characterizing traffic conditions in urban areas. *Trans. Sci.* 18.
- Herman, R. and I. Prigogine (1979). A two-fluid approach to town traffic. *Science* 204.
- Lighthill, M. J. and G. B. Whitham (1955). On kinematic waves II: A theory of traffic flow on long crowded roads. *Pro. of the Royal Soc. A* 229.
- Mahmassani, H. S., Jayakrishnan, R. and R. Herman (1990). Network traffic flow theory: Microscopic simulation experiments on supercomputers. *Trans. Res. A* 24.
- Munakata, K. (2007). Examination of correlation between travel times on links. *ME Res. Report*, Univ. of Canterbury.
- Nicholson, A. and K. Munakata (2009). Estimating the benefits of trip time reliability. *32<sup>nd</sup> Australasian Trans. Res. Forum*.
- Nicholson, A., Kuhn, K. and K. Munakata (2010). Travel time reliability and correlation. submitted to *Trans. Res. A*.
- Olszewski, P., H. S. L. Fan and Y.-W. Tan (1995). Area-wide traffic speed-flow model for the Singapore CBD. *Trans. Res. A* 29.
- Papageorgiou, M. (1998). Some remarks on macroscopic traffic flow modelling. *Trans. Res. A* 32.
- Shin, J.-S. (1987). Aggregation and the endogeneity problem. *Int. Econ. Journal* 1.
- Thiel, H. (1954). *Linear aggregation of economic relations*. North-Holland, Amsterdam.
- Thomson, J. M. (1967). Speeds and flow of traffic in central London: Parts I and II. *Traf. Eng. and Control* 8.
- Wardrop, J. G. (1968). Journey speed and flow in central urban areas. *Traf. Eng. and Control* 9.
- Williams, J. C., Mahmassani, H. S., and R. Herman (1987). Urban traffic network flow models. *Trans. Res. Rec.* 1112.
- Yule, G. (1903). Notes on the theory of associations of attributes in statistics. *Biometrika* 2.